ELSEVIER Stochastic Processes and their Applications 54 (1994) 291–296

stochastic processes and their applications

On Stein's method and point process approximation

Timothy C. Brown, Darryl Greig*

Department of Statistics, The University of Melbourne, Parkville 3052, Australia

Received 25 July 1993; revised 10 November 1992

Abstract

A correction to an error in a recent paper of Barbour and Brown's is detailed, and a relevant corollary stated.

Keywords: Poisson approximation; Hard core process

1. Introduction

Stein's method has been applied by Barbour and Brown (1992) to derive the total variation distance between a finite point process and a Poisson process. They apply the results obtained to the hard core point process over the d-dimensional torus $[0,1]^d$, with hard core radius r: this process is Poisson but for the restriction that no two points in the process are distance less than r apart. The computation of the stated upper bound (Barbour and Brown, 1992, Corollary 2.5) is not correct due to an error in the determination of the third term of Eq. (2.8) of the same paper.

In this note we propose a correction to the above example, and suggest an alternative method using a corollary to Theorem 2.4 of Barbour and Brown (1992) which is simpler and gives better bounds in some cases.

2. Preliminaries

Let Γ be a compact, second countable Hausdorff space with some fixed measure ν, and let $\mathscr{H}$ denote the space of *finite* point process configurations on Γ. Suppose Ξ is a simple point process on Γ for which Janossi densities $j_n: \Gamma^n \to \mathbb{R}^+$ with respect to ν^n exist for all $n \in \mathbb{Z}^+$. That is, for any non-negative measurable function $f: \mathscr{H} \to \mathbb{R}$,

$$\boldsymbol{E}(f(\Xi)) = \sum_{n \geq 0} \int_{\Gamma^n} (n!)^{-1} f\left(\sum_{i=1}^{n} \delta_{\alpha_i}\right) j_n(\alpha_1, \ldots, \alpha_n) \nu^n(\mathrm{d}\alpha_1, \ldots, \mathrm{d}\alpha_n). \tag{1}$$

* Corresponding author.

SSDI 0304-4149(94)00018-O

Also suppose that Ξ has finite mean measure $\boldsymbol{\pi}$ and let μ be the density with respect to ν of $\boldsymbol{\pi}$. Then we may write

$$\mu(\alpha) = \sum_{n \geq 0} \int_{\Gamma^n} (n!)^{-1} j_{n+1}(\alpha, \alpha_1, \ldots, \alpha_n) \nu^n(\mathrm{d}\alpha_1, \ldots, \mathrm{d}\alpha_n). \tag{2}$$

Suppose that for each $\alpha \in \Gamma$ there is some neighbourhood N_α containing α such that the mappings

$$\mathscr{H} \times \Gamma \to [0, \infty) : (\xi, \alpha) \mapsto \xi(N_\alpha)$$

and

$$\mathscr{H} \times \Gamma \to \mathscr{H} : (\xi, \alpha) \mapsto \xi \text{ restricted to } N_\alpha$$

are product measurable. Let Ξ^α denote the configuration of Ξ outside of N_α. Then for some $\alpha \in \Gamma$ let m be fixed and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_m) \in (N_\alpha^c)^m$ and define

$$g(\alpha, \boldsymbol{\beta}) = \frac{\sum_{r \geq 0} \int_{N_\alpha^r} j_{m+r+1}(\alpha, \boldsymbol{\beta}, \boldsymbol{\gamma}) (r!)^{-1} \nu^r(\mathrm{d}\boldsymbol{\gamma})}{\sum_{s \geq 0} \int_{N_\alpha^s} j_{m+s}(\boldsymbol{\beta}, \boldsymbol{\eta}) (s!)^{-1} \nu^s(\mathrm{d}\boldsymbol{\eta})}. \tag{3}$$

Note that for any $\boldsymbol{\beta}$ which is a legitimate configuration of Ξ^α, (3) gives the conditional density of a point at α given the configuration of Ξ outside of N_α.

If $\boldsymbol{P}$ and $\boldsymbol{Q}$ are two probability measures on $\mathbb{N}$, then the *total variation distance*, d_{TV}, between them is defined by

$$d_{\mathrm{TV}}(\boldsymbol{P}, \boldsymbol{Q}) = \frac{1}{2} \sum_{n=0}^{\infty} |\boldsymbol{P}(n) - \boldsymbol{Q}(n)| = \sup_{A \subset \mathbb{N}} |\boldsymbol{P}(A) - \boldsymbol{Q}(A)|.$$

For any random variable X, denote its probability law by $\mathscr{L}X$.

Now Theorem 2.4 of Barbour and Brown (1992) may be stated.

Theorem 1. *Under the conditions above and for any finite measure $\boldsymbol{\lambda}$ on Γ,*

$$\begin{aligned} d_{\mathrm{TV}}(\mathscr{L}(\Xi), \mathrm{Poisson}(\boldsymbol{\lambda})) \leq \boldsymbol{E}\left\{\int_\Gamma \Xi(N_\alpha \setminus \{\alpha\}) \Xi(\mathrm{d}\alpha)\right\} + \int_\Gamma \boldsymbol{E}\{\Xi(N_\alpha)\} \mu(\alpha) \nu(\mathrm{d}\alpha) \\ + \int_\Gamma \boldsymbol{E}|g(\alpha, \Xi^\alpha) - \mu(\alpha)| \nu(\mathrm{d}\alpha) + \|\boldsymbol{\pi} - \boldsymbol{\lambda}\|, \end{aligned} \tag{4}$$

where $\|\cdot\|$ denotes the total variation of the signed measure of $\boldsymbol{\pi} - \boldsymbol{\lambda}$.

Corollary 1. *Let $\Xi^{\{\alpha\}}$ be the configuration of Ξ outside of the singleton $\{\alpha\}$, then,*

$$\begin{aligned} d_{\mathrm{TV}}(\mathscr{L}(\Xi), \mathrm{Poisson}(\boldsymbol{\lambda})) \leq \int_\Gamma \boldsymbol{E}\{\Xi(\{\alpha\})\} \mu(\alpha) \nu(\mathrm{d}\alpha) \\ + \int_\Gamma \boldsymbol{E}|g(\alpha, \Xi^{\{\alpha\}}) - \mu(\alpha)| \nu(\mathrm{d}\alpha) + \|\boldsymbol{\pi} - \boldsymbol{\lambda}\|. \end{aligned} \tag{5}$$

Proof. Take $N_\alpha = \{\alpha\}$ in Theorem 1. □

The definitions may now be specialised to the example at hand. Let Γ be the d-dimensional torus $[0,1]^d$, and let κ_d be the volume of the unit ball in d dimensions. Let Ξ be the hard core point process with hard core radius r and total number of points λ. The process Ξ may be specified by its Janossi density with respect to Lebesgue measure ν on Γ, that is

$$j_n(\boldsymbol{\alpha}) = c\kappa^n \boldsymbol{I}\left[\bigcap_{i \neq j} |\alpha_i - \alpha_j| > r\right],$$

where c, the partition function of statistical physics, and κ are numbers depending on λ and r chosen in such a way that

$$\sum_{n \geq 0} \int_{\Gamma^n} (n!)^{-1} j_n(\boldsymbol{\alpha}) \nu^n(\mathrm{d}\boldsymbol{\alpha}) = 1 \quad \text{and} \quad \int_{\Gamma} \mu(\alpha)\nu(\mathrm{d}\alpha) = \lambda.$$

Note that, in this example, $\mu(\alpha) = \lambda$ for all $\alpha \in \Gamma$.

The error in Barbour and Brown (1992) occurs in the computation of (3) for this process. Namely, when one considers the denominator of $g(\alpha, \boldsymbol{\beta})$, for some $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_m\} \in (N_\alpha^c)^m$, there can be more terms than the stated $s = 0$ and $s = 1$. For example if no points of $\boldsymbol{\beta}$ lie in the ball $B_\alpha(2r)$ there may be many points in $N_\alpha = B_\alpha(r)$ (up to 2 for dimension $d = 1$, up to 5 for $d = 2$ and more for $d > 2$), each one distance $\geq r$ away from each other and the points of $\boldsymbol{\beta}$. This problem is rectified in the following sections by considering neighbourhoods of radius $\leq r/2$ in which at most one point of a legitimate configuration of the hard core process can occur.

3. An upper bound for the hard core process

In the first instance we apply the result obtained in Corollary 1, which is equivalent to using a neighbourhood of radius zero around each $\alpha \in \Gamma$. It turns out that the evaluation of the upper bound in this corollary for the hard core process is essentially exact, although in Section 4 we will demonstrate that under a further condition on the parameters of the process, a better bound is obtainable from Theorem 1 using a neighbourhood of non-zero radius.

To begin with, note that the first term of the right-hand side of (5) is 0 because $\boldsymbol{E}\{\Xi(\{\alpha\})\} = 0$. Further, since ν in (3) is Lebesgue measure for this example, and $\nu(N_\alpha) = \nu(\{\alpha\}) = 0$, the numerator of $g(\alpha, \Xi^{\{\alpha\}})$ only involves the term $r = 0$, and the denominator $s = 0$. Let $m \geq 1$ be fixed and let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_m)$ be some legitimate configuration of $\Xi^{\{\alpha\}}$. That is $\alpha \notin \boldsymbol{\beta}$ and $|\beta_i - \beta_j| > r \ \forall i \neq j$. Then

$$g(\alpha, \boldsymbol{\beta}) = \frac{j_{m+1}(\alpha, \boldsymbol{\beta})}{j_m(\boldsymbol{\beta})} = \frac{c\kappa^{m+1} \boldsymbol{I}[\bigcap_i |\alpha - \beta_i| > r]}{c\kappa^m}$$

$$= \kappa \boldsymbol{I}\left[\bigcap_i |\alpha - \beta_i| > r\right] = \kappa \boldsymbol{I}\left[\sum_{i=1}^m \delta_{\beta_i}(B_\alpha(r)) = 0\right]$$

and so

$$g(\alpha, \Xi^{\{\alpha\}}) = \kappa \boldsymbol{I}[\Xi^{\{\alpha\}}(B_\alpha(r)) = 0]. \tag{6}$$

Thus since $\mu(\alpha) = g(\alpha, \Xi^{\{\alpha\}})$ takes the value λ everywhere in this example, it follows from (6) that

$$\boldsymbol{P}(\Xi^{\{\alpha\}}(B_\alpha(r)) = 0) = \frac{\lambda}{\kappa}. \tag{7}$$

Then the expectation in the second term of (5) may be bounded above in the following way:

$$\begin{aligned}
\boldsymbol{E}|g(\alpha, \Xi^{\{\alpha\}}) - \mu(\alpha)| &= \boldsymbol{E}|\kappa \boldsymbol{I}[\Xi^{\{\alpha\}}(B_\alpha(r)) = 0] - \lambda| \\
&= (\kappa - \lambda)\boldsymbol{P}(\Xi^{\{\alpha\}}(B_\alpha(r)) = 0) + \lambda \boldsymbol{P}(\Xi^{\{\alpha\}}(B_\alpha(r)) \geq 1) \\
&= (\kappa - \lambda)\frac{\lambda}{\kappa} + \lambda\left(1 - \frac{\lambda}{\kappa}\right) = 2\lambda \boldsymbol{P}(\Xi^{\{\alpha\}}(B_\alpha(r)) \geq 1) \\
&\leq 2\lambda \boldsymbol{E}[\Xi(B_\alpha(r))] = 2\lambda^2 \kappa_d r^d.
\end{aligned}$$

Now it is immediate from Corollary 1 that

$$d_{TV}(\mathscr{L}(\Xi), \text{Poisson}(\boldsymbol{\pi})) \leq 2\lambda^2 \kappa_d r^d. \tag{8}$$

4. A tighter bound for the hard core process

Despite the apparent exactness of the computation in Section 3, it turns out that by applying Theorem 1 directly under more stringent conditions on λ, κ and r, we may improve the constant multiplier of (8). The derivation used by Barbour and Brown (1992) is closely followed, but instead of the neighbourhood $N_\alpha = B_\alpha(r)$ in Theorem 1, the neighbourhood $N_\alpha = B_\alpha(r/n)$, for some $n \in \mathbb{Z}^+$ such that $n \geq 2$ is used. In this case it is clear that there can be at most one point occurring in N_α for any legitimate configuration of Ξ.

The first two terms of the right-hand side of (4) follow through in precisely the original manner with the new neighbourhood. That is, the first term reduces to 0 for all $n \in \mathbb{Z}^+$, and the second term may be simply computed from the Janossi densities of the process to be $\lambda^2 \kappa_d r^d n^{-d}$.

Some subtleties are introduced into the computation of the third term by the new N_α. We begin by considering (3) where $m \geq 1$ is fixed and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_m)$ is some legitimate configuration of Ξ^α. That is $\alpha \notin \boldsymbol{\beta}$ and $|\beta_i - \beta_j| > r \ \forall i \neq j$.

Clearly the numerator can only involve the term with $r = 0$, and thus reduces to $j_{m+1}(\alpha, \boldsymbol{\beta})$. However note that this may also be zero if any $\beta_i \in \boldsymbol{\beta}$ is within r of α. Hence the numerator may be written as

$$c\kappa^{m+1} \boldsymbol{I}\left[\sum_{i=1}^{m} \delta_{\beta_i}(B_\alpha(r)) = 0\right].$$

In the denominator the only possible terms are $s = 0$ and $s = 1$, and so the denominator is $c\kappa^m\{1 + \kappa v(A_\alpha(\boldsymbol{\beta}))\}$, where $A_\alpha(\boldsymbol{\beta}) = \{\alpha' \in N_\alpha : |\beta_i - \alpha'| > r \ \forall i\}$. Then $0 \leq v(A_\alpha(\boldsymbol{\beta})) \leq \kappa_d(r/n)^d$. Thus

$$\frac{\kappa \boldsymbol{I}[\sum_{i=1}^{m} \delta_{\beta_i}(B_\alpha(r)) = 0]}{\{1 + \kappa\kappa_d(r/n)^d\}} \leq g(\alpha, \boldsymbol{\beta}) \leq \kappa \boldsymbol{I}\left[\sum_{i=1}^{m} \delta_{\beta_i}(B_\alpha(r)) = 0\right], \tag{9}$$

so we may write $g(\alpha,\boldsymbol{\beta})=0$ iff $\sum_{i=1}^{m}\delta_{\beta_i}(B_\alpha(r))\geq 1$. Also if $\sum_{i=1}^{m}\delta_{\beta_i}(B_\alpha(r))=0$ then $\kappa\{1+\kappa\kappa_d(r/n)^d\}^{-1}\leq g(\alpha,\boldsymbol{\beta})\leq\kappa$.

Now the required probability may be bounded above by

$$\begin{aligned}\boldsymbol{P}[\Xi^\alpha(B_\alpha(r))\geq 1] &= \boldsymbol{P}\left[\left(\Xi(B_\alpha(r))-\Xi\left(B_\alpha\left(\frac{r}{n}\right)\right)\right)\geq 1\right]\\ &\leq \boldsymbol{E}\left[\Xi(B_\alpha(r))-\Xi\left(B_\alpha\left(\frac{r}{n}\right)\right)\right]\\ &= \lambda\kappa_d r^d-\lambda\kappa_d r^d n^{-d},\end{aligned}\tag{10}$$

so

$$\begin{aligned}\lambda=\mu(\alpha)=\boldsymbol{E}g(\alpha,\Xi^\alpha) &\geq \frac{\kappa}{\{1+\kappa\kappa_d(r/n)^d\}}\,\boldsymbol{P}[\Xi^\alpha(B_\alpha(r))=0]\\ &\geq \frac{\kappa}{\{1+\kappa\kappa_d(r/n)^d\}}(1-\lambda\kappa_d r^d+\lambda\kappa_d r^d n^{-d}).\end{aligned}$$

By a simple rearrangement and noting the upper bound on λ given by (9) it follows that

$$\lambda\leq\kappa\leq\frac{\lambda}{1-\lambda\kappa_d r^d},\tag{11}$$

where $\lambda\kappa_d r^d$ is assumed to be <1.

Let $n_0=((2\lambda-\kappa)\kappa\kappa_d r^d/(2\kappa-2\lambda))^{1/d}$. If it is known that $\Xi^\alpha(B_\alpha(r))=0$, then

$$\begin{aligned}|g(\alpha,\Xi^\alpha)-\mu(\alpha)| &\leq \max\left\{(\kappa-\lambda),\left(\lambda-\frac{\kappa}{1+\kappa\kappa_d r^d n^{-d}}\right)\right\}\\ &= \begin{cases}(\kappa-\lambda) & \text{if } n>n_0,\\ \left(\lambda-\dfrac{\kappa}{1+\kappa\kappa_d r^d n^{-d}}\right) & \text{if } n\leq n_0.\end{cases}\end{aligned}$$

Now, for $n\leq n_0$, the expected difference in the third term is bounded in the following way:

$$\begin{aligned}&\boldsymbol{E}|g(\alpha,\Xi^\alpha)-\mu(\alpha)|\\ &\leq\left(\lambda-\frac{\kappa}{1+\kappa\kappa_d r^d n^{-d}}\right)\boldsymbol{P}(\Xi^\alpha(B_\alpha(r))=0)+\lambda\,\boldsymbol{P}(\Xi^\alpha(B_\alpha(r))\geq 1)\\ &\leq\lambda-\frac{\kappa}{1+\kappa\kappa_d r^d n^{-d}}(1-\lambda\kappa_d r^d+\lambda\kappa_d r^d n^{-d})\\ &=\frac{\lambda-\kappa+\lambda\kappa\kappa_d r^d}{1+\kappa\kappa_d r^d n^{-d}}\leq\frac{\lambda^2\kappa_d r^d}{1+\lambda\kappa_d r^d n^{-d}},\end{aligned}$$

where the second inequality follows from (10), and the third from (11).

Thus, by adding the second and third terms we may deduce from Theorem 1 that, for $\lambda\kappa_d r^d < 1$ and $2 \le n \le n_0$,

$$d_{\mathrm{TV}}(\mathscr{L}(\Xi), \mathrm{Poisson}(\boldsymbol{\pi})) \le \lambda^2 \kappa_d r^d \left(\frac{1}{1 + \lambda\kappa_d (r/n)^d} + n^{-d} \right).$$

This bound achieves its minimum value at $n = 2$, and hence

$$d_{\mathrm{TV}}(\mathscr{L}(\Xi), \mathrm{Poisson}(\boldsymbol{\pi})) \le \lambda^2 \kappa_d r^d \left(\frac{1}{1 + \lambda\kappa_d r^d 2^{-d}} + 2^{-d} \right) \le (1.5)\lambda^2 \kappa_d r^d \tag{12}$$

and note that as $d \to \infty$, the total variation distance converges to $\lambda^2 \kappa_d r^d$.

By a similar computation in the $n > n_0$ case the following bound is obtained:

$$d_{\mathrm{TV}}(\mathscr{L}(\Xi), \mathrm{Poisson}(\boldsymbol{\pi})) \le 2\lambda^2 \kappa_d r^d + \frac{\lambda^3 (\kappa_d r^d)^2 n^{-d}}{1 - \lambda\kappa_d r^d}, \tag{13}$$

which converges to (8) as $n \to \infty$, but is greater for all finite n.

5. Concluding remarks

Since the lower bound obtained in Corollary 2.5 of Barbour and Brown (1992) is correct as stated, both the upper bounds in (8) and (12) are tight as $\lambda r^d \to 0$. Note that if $\lambda\kappa_d r^d \le \frac{1}{2}$, the ratio to first order of the upper to lower bounds of (8) and (12) are 4 and 3, respectively. Clearly the better constant factor of (12) makes this the tighter of the bounds, however it is not at all clear that the condition $n_0 \ge 2$ will be satisfied in cases of interest, and thus whether this bound may be applied in general. One would expect a more detailed analysis of the relationship between κ, λ and r to provide the circumstances under which (12) is appropriate.

In this particular example the application of Corollary 1 gives a very simple upper bound calculation that appears to be better than that obtainable by a direct application of Theorem 1 in some cases, although it is not clear that this would generally be true. Certainly there is an intuitive appeal to this corollary since $\boldsymbol{E}|g(\alpha, \Xi^{\{x\}}) - \mu(\alpha)|$ may be thought of as a measure of the deviance away from Poisson of a process in terms of the dependence between the density of a point α and the conditional density of α given the influence of its immediate neighbours.

Finally note that the above corrections should also be applied to the example following Theorem 3.6 of Barbour and Brown (1992).

Acknowledgements

The authors are grateful to the referee for his comments that improved the article, and to Aihua Xia for his substantial contribution towards identifying the problem.

References

A.D. Barbour and T.C. Brown, Stein's method and point process approximation, Stochastic Process. Appl. 43 (1992) 9–31.